DOCUMENT RESUME

ED 075 273 SO 005 425

AUTHOR	Harris, Margaret L.; Tabachnick, B. Robert
TITLE	Measuring Social Studies Concept Attainment: Boys and Girls. Report from the Project on A Structure of Concept Attainment Abilities.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Research and Development Centers Branch.
REPORT NO	TR-193
BUREAU NO	BR-5-0216
PUB DATE	Nov 71
CONTRACT	OEC-5-10-154
NOTE	34p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Achievement Tests; Cognitive Measurement; *Concept Formation; Educational Research; Elementary School Students; Geography; Grade 5; Grade 6; *Measurement Instruments; *Measurement Techniques; Performance Criteria; *Social Studies; *Test Construction; Testing Problems

ABSTRACT

This paper describes test development efforts for measuring achievement of selected concepts in social studies. It includes descriptive item and test statistics for the tests developed. Twelve items were developed for each of 30 concepts. Subject specialists categorized the concepts into three major areas: Geographic Region, Man and Society, and Map and Globe Study. The items were administered during early summer of 1970 to 196 girls who had just completed the fifth grade and during the fall of 1970 to 195 boys who had just begun the sixth grade. Some of the major conclusions are: 1) reliability estimates obtained for both total concept scores and total task scores are sufficiently high to warrant study of the dimensionality of these social studies concepts and the dimensionality of the tasks when using social studies content; 2) the three area distinctions seem not to be important ones; 3) the difficulty item indices obtained indicate that these items are of appropriate difficulty levels for these subjects; and 4) almost all the items have desirable levels of discrimination indices when the item is both a part of a concept criterion score and a task criterion score. Related reports from the Project are ED 065 894 and ED 068 410. (FDI)

FILMED FROM BEST AVAILABLE COPY

EU 002415

BR 5-0216
50

TECHNICAL REPORT NO. 193

MEASURING SOCIAL STUDIES CONCEPT ATTAINMENT: BOYS AND GIRLS

REPORT FROM THE PROJECT ON A STRUCTURE OF CONCEPT ATTAINMENT ABILITIES

SD 005425

WISCONSIN RESEARCH AND DEVELOPMENT

CENTER FOR COGNITIVE LEARNING

THE UNIVERSITY OF WISCONSIN
Madison, Wisconsin

U. S. Office of Education
Center No. C-03
Contract OE 5-10-154

RESEARCH
COGNITIVE LEARNING
DEVELOPMENT

ED 075273

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Technical Report No. 193

MEASURING SOCIAL STUDIES CONCEPT ATTAINMENT:

BOYS AND GIRLS

By Margaret L. Harris and B. Robert Tabachnick

Report from the Project on
A Structure of Concept Attainment Abilities

Robert E. Davidson, Lester S. Golub, Herbert J. Klausmeier,
Thomas A. Romberg, B. Robert Tabachnick, Alan M. Voelker
Principal Investigators

The Quality Verification Program
Mary R. Quilling, Director

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

November 1971

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

Statement of Focus

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Quality Verification Program and from the Project on the Structure of Concept Attainment Abilities in Program 1. The Quality Verification Program assisted in developing tests to measure concept achievement and identifying reference tests for cognitive abilities, while the Concept Attainment staff took primary initiative in identifying basic concepts in social studies at intermediate grade level. The tests will be used to study the relationships among cognitive abilities and learned concepts in various subject matter areas. The outcome of the Project will be a formulation of a model of structure of abilities in concept attainment in a number of subjects, including mathematics, science, and language arts, as well as social studies.

Contents

		Page
	List of Tables and Figures	vii
	Abstract	ix
I	Introduction	1
II	Procedures	5
	Test Development	5
	Readability	5
	Validity	5
	Reliability	6
	Item Revision	7
	Subjects	8
	Data Collection	8
	Treatment of the Data	9
III	Results and Discussion	13
	Reliability Estimates and Test Statistics	13
	Item Indices	16
IV	Summary and Conclusions	27
	Conclusions	27
	Recommendation	27
	References	29
	Appendix A: Key for Concepts in Numerical Order	31

List of Tables

Table		Page
1	Social Studies Concepts Categorized by Area	2
2	Test Data for Population and Samples	9
3	Distribution of Fathers' Occupations	10
4	Means, Standard Deviations, and Reliability Estimates for Social Studies Concept Scores: Boys and Girls	14
5	Means, Standard Deviations, and Reliability Estimates for Social Studies Task Scores: Boys and Girls	15
6	Reliability Estimates for Task Scores by Area and Total for Girls	16
7	Item Indices Based on Concept and Task Criterion Scores	17

List of Figures

Figure		Page
1	Item matrix for each individual.	6

[1]Tables 4, 5, and 7 were prepared from the computer output by Mr. Gert Huysamen and are being used in his Doctoral Dissertation.

Abstract

Test development efforts for constructing 12 items to measure achievement of each of 30 selected social studies concepts are described. Item and total score statistics for data collected on 196 girls who had just completed the fifth grade and 195 boys who had just begun the sixth grade are presented and discussed.

I
Introduction

The primary objective of the project entitled "A Structure of Concept Attainment Abilities" (hereafter referred to as the CAA Project) is to formulate one or more models or structures of concept attainment abilities, and to assess their consistency with actual data. The major steps for attaining this primary objective were taken to be:

1. To identify basic concepts in language arts, mathematics, science, and social studies appropriate at the fourth grade level,
2. To develop tests to measure achievement of these concepts,
3. To identify reference tests for cognitive abilities, and
4. To study the relationships among learned concepts in these four subject matter fields and the identified cognitive abilities.

This paper describes the test development efforts for measuring achievement of selected concepts in social studies; thus, it is a report of one aspect of Step 2. As such, it will include descriptive item and test statistics for the tests developed. The items can be found in "Items to Test Level of Attainment of Social Studies Concepts by Intermediate Grade Children" (Tabachnick, Weible, & Livermore, 1970).

Concepts may be defined in one or more of four ways: (a) structurally, in terms of perceptible or readily specifiable properties or attributes; (b) semantically, in terms of synonyms or antonyms; (c) operationally, in terms of the procedures employed to distinguish the concept from other concepts; or (d) axiomatically, in terms of logical or numerical relationships (Klausmeier, Harris, Davis, Schwenn, & Frayer, 1968). "A concept exists whenever two or more distinguishable objects or events have been grouped or classified together and set apart from objects on the basis of some common feature or property of each" (Bourne, 1966, p. 1). The concept of Bourne's definition might be called a classificatory one and seems to be the same as the structural type discussed by Klausmeier et al. (1968). This is the type of concept with which this project is concerned, and such a definition of a concept served as the basis for selection and analysis of subject matter concepts.

Many different types of performance might be taken as the critical evidence that a student does or does not understand a given concept. Thus, as a part of this project it is necessary to have a schema for measuring understanding of concepts. Such a schema was developed by Frayer, Fredrick, and Klausmeier (1969) and was used by the CAA Project to assess concept attainment. The "Schema for Testing the Level of Concept Mastery" consists of 13 types of questions, each involving a different task required of the examinee. The schema also allows for selection of an answer (multiple-choice type questions) or for production of an answer (completion type questions). It was decided to use the first 12 tasks and a multiple-choice format for this project. The 12 tasks of the schema which were used are:

1. Given the name of an attribute, select an example of the attribute.
2. Given an example of an attribute, select the name of the attribute.
3. Given the name of a concept, select an example of the concept.
4. Given the name of a concept, select a nonexample of the concept.
5. Given an example of a concept, select the name of the concept.
6. Given the name of a concept, select the relevant attribute.

Table 1. Social Studies Concepts Categorized by Area

Geographic Region	Man and Society	Map and Globe Study
Bay	Agriculture	Area (square miles)
Canal	*Airway	Axis
Climate	Basic Needs	Boundary
*Coastline	*City	Continent
*Delta	Commerce (trade)	*Country
*Desert	*Countryside	Day
Elevation	*Democracy	*Distance
Geography	Economy	Earth
*Gulf	Educational Institution	*East-West (lines of latitude)
Harbor	*Exchange	Equator
Highland	Family	*Globe
Hills	Farming	Gravity
Island	Fishing	Hemispheres
Isthmus	Forestry	Legend
Lake	*Government	Map
Location	Industry	*Map Directions
Mountain	Institutions	*Map Measurement
Mountain Pass	International	*Map Scale
Mountain Peak	*Land Routes	Meridians
*Mountain Region	Man	Model
Ocean Currents	Man as a member of a group	Night
Ocean Tides	Man as an individual	*North-South (lines of longitude)
Peninsula	Manufacturing	Ocean
Plain	Market	Orientation
Prairie	Nation	Parallels
Precipitation	Nature	*Physical Feature Map
Region	*News	Planet
Ridge	*Organization	Political Map (of nations, countries)
*River	President	Revolution
*River Mouth	Republic	Rotation
River Source	Service Organization	Sea Level-Below Sea Level
*Strait	Society	Seasons
Subtropical Region	State	Solar System
Swamp	Suburban	*Symbol Map
Temperature	Transportation	Topographical Map (map of land forms)
Topography	Urban	
Transitional Region	Village	
*Tributary	*Waterway	
*Tropical Region		
Valley		
Waterway		
Weather		

*Concepts randomly selected to be tested.

7. Given the name of a concept, select the irrelevant attribute.
8. Given the definition of a concept, select the name of the concept.
9. Given the name of a concept, select the definition of the concept.
10. Given the name of a concept, select the supraordinate concept.
11. Given the name of a concept, select the subordinate concept.
12. Given the name of two concepts, select the relationship between them.

Single- or compound-word classificatory concepts (those that are defined by attributes) in social studies subject matter at the fourth grade level were identified. This task was subdivided into four steps:

1. Identification of the major areas within the subject matter of social studies,
2. Selection of three of these major areas to be studied,
3. Identification of classificatory concepts within each of these three major areas, and
4. Random sampling of ten concepts from those identified for each of the three major selected areas.

This yielded a total of 30 social studies concepts to be studied by the project. A list is given in Table 1, by area, of the concepts identified and randomly selected for study. The areas are Geographic Region, Man and Society, and Map and Globe Study. A description of the procedures used to identify these concepts can be found in "Selection and Analysis of Social Studies Concepts for Inclusion in Tests of Concept Attainment" (Tabachnick, Weible, & Frayer, 1970).

The researchers of Project 101, Situational Variables and Efficiency of Concept Learning, developed a system for analyzing a concept in preparation for developing items to measure the level of attainment of that concept (Frayer, Fredrick, & Klausmeier, 1969). Since the publication of that paper they, in cooperation with the researchers of the CAA Project, have refined their thinking and advanced this system. The refinements are discussed in "A Structure of Concept Attainment Abilities: The Problem and Strategies for Attacking It" (Harris, Harris, Frayer, & Quilling, in press). Briefly, a concept may be described in many ways: in terms of its criterial, relevant, and irrelevant attributes; its examples and nonexamples; its supraordinate, coordinate, and subordinate hierarchical relationships (theoretically determined); and its lawful or other types of relationships to other concepts. Knowledge of each of these kinds of information may be tested to determine a student's level of attainment of a concept. An analysis, along these lines, of each of the 30 sampled social studies concepts which are being studied can be found in "Selection and Analysis of Social Studies Concepts for Inclusion in Tests of Concept Attainment" (Tabachnick, Weible, & Frayer, 1970).

Thus, using the analysis of a concept as the basis for appropriate content and the 12 tasks of the schema as the basis for appropriate tasks, 12 items were developed for each of the 30 concepts. There was one item for each of the 12 tasks (except for Concept 10 which had no Task 11), making a total of 359 social studies items which were developed for the purpose of measuring and assessing concept attainment in social studies. The development of the items, along with item and total score statistics (for concepts and for tasks) obtained for them for fifth grade boys and girls, will be discussed in the following sections.

II
Procedures

This section contains a discussion of the item development procedures used including initial item construction and revision of those items based on item analysis results. Also included is a discussion of the data collection procedures, subjects, and treatment of the data.

Test Development

One item for each of the 12 tasks was generated for each of the 30 selected concepts, with the exception of Concept 10, Task 11. If one looks at the tasks used to measure understanding of the concept, it is apparent that there can be more than one item generated for at least some of the tasks. For example, a Task 1 type item could be constructed to measure understanding of each of many relevant attributes for most concepts. For this project, it was decided to construct just one multiple-choice item for each task for each concept. This made it necessary to have bases for making choices when such choices were necessary. These bases consisted of principles for selecting attributes, relationships, incorrect choices, etc. A discussion of such bases may be found in "A Structure of Concept Attainment Abilities: The Problem and Strategies for Attacking It" (Harris et al., in press).

General procedures for item construction included initial item generation by a subject matter specialist item writer; critique of the items by a committee composed of the item writers from each of the four subject matters being studied (the other three are language arts, mathematics, and science), an experienced elementary school teacher specializing in reading, and a measurement specialist; and final critique by the subject matter principal investigator and a measurement specialist. Concerns in the item construction process were readability, validity, and reliability.

Readability

It was intended that no student should be unable to answer an item correctly simply because of inability to read the item. In writing items, very simple language was used wherever possible. Several pilot studies concerned with the readability question were conducted, and two outside consultants expert in the testing and measurement fields were asked to look at a sample of the items from the point of view of readability for fifth graders. No significant differences were found among treatment groups; percentage of occurrence of subjects who could not pronounce the word and did not know its meaning when shown the concept labels, but did know its meaning when the word was pronounced, was judged to be negligible; and the two outside consultants independently advised that there was no reading problem with the items and that there should be no concern about administering them in the standard way in which the students read the items themselves. The conclusion drawn from the results of the pilot studies and the consultants' opinions was that readability of the items was not a problem and standard administration conditions would be satisfactory. For further information see Harris et al. (in press).

Validity

The content validity of each of the items was of immediate concern during item construction; aspects of construct validity were to be probed later using duplicate test construction, simplex analysis, and factor analysis of the results obtained using the content-valid items constructed.

Content Validity. Each item was constructed to meet the content and task specifications set for it. The task required of the student by each item was specified by the

schema adopted for use in measuring concept attainment. The concept name was given by the sampling process; the attributes, examples, definition, and relationships associated with the concept name were defined by the prior analysis of the concept. The content for each item was specified in this manner. The content specifications were not as precise as the task specifications due to the necessity of choosing a single attribute to be tested for example and selecting the incorrect alternatives to be used in the multiple-choice questions. Systematic construction of alternate choices was used whenever possible; for example, for an item dealing with a path of travel, other types of travel (or examples of them) were used as incorrect choices, e.g. airway, waterway, land route.

To further ensure the content validity of the items, two persons who were familiar with the schema for testing concept attainment, but were not involved in the item development process, classified five random sets of 72 items (12 items for six concepts in each set), according to content and task. These two persons had the analysis of the concepts available. They were able to correctly classify all but a few of the items. Any questions they had about these few items were mutually resolved among the subject matter principal investigator, the measurement specialist, and themselves.

Reliability

Developing one item for each of the 12 tasks for each of the 30 selected concepts yields a 12 (tasks) by 30 (concepts) matrix consisting of the score for each of the 360 items, one for each cell of the matrix, for each individual to whom the items are administered. Thus, a completely crossed design exists and two types of total scores can be secured from this matrix: a total score for each of the 30 concepts (totalled across tasks) and a total score for each of the 12 tasks (totalled across concepts). Figure 1 is an illustration of such a matrix.

	CONCEPTS																														
	Area 1										Area 2										Area 3										Total Score for Tasks
	1	2								10	11	12								20	21	22								30	
1																															
2																															
TASKS .																															
12																															
Total Score for Concepts																															

Fig. 1. Item matrix for each individual.

This design offers these alternatives: (a) use a total score of 360 items to analyze all items against; (b) use 30 total scores, each for one concept and consisting of 12 items, to analyze the 12 task items against; and (c) use 12 total scores, each for one task and consisting of 30 items, to analyze the 30 concept items against. The first alternative was rejected since it assumes neither task nor concept variation is present. A choice was not made between the next two alternatives. Instead, both were done. An important theoretical problem of how to item analyze a completely crossed design like this remains to be solved.

Major concerns about reliability for the test development process were that internal consistency reliability estimates for task scores (total of 30 items across concepts) and concept scores (total of 12 items across tasks) be high enough to warrant further study using such scores. It was recognized that there might be some contradictions in what was attempted. The items were constructed to comply with the completely crossed design, 30 concepts by 12 tasks. One major objective of the entire project is to determine the dimensionality of the selected social studies concepts and of the tasks when using social studies content. If either or both of these are not unidimensional, then an internal consistency reliability estimate based upon items measuring aspects from the multidimensions would reflect this; the more dimensions present and the more uncorrelated they are, the lower the internal consistency estimate. Recognizing this, and not being able to study the dimensionality of the two modes (concepts and tasks) until after the items were developed, pilot studies were conducted using the items for some of the concepts for the 12 tasks. As will be pointed out later, evidence indicates that sufficiently reliable scores can be obtained for both task scores and concept scores.

Item Revision

If one looks at the 12 tasks for a single concept it becomes quite apparent that there may be a strong learning effect as one attempts to answer the items. The name of the concept appears in every item, except for the first two which deal with an attribute of the concept, either in the stem or as a possible choice. This makes a random presentation of the items desirable. Using items for six of the mathematics concepts presented on mark sense type cards, a study was conducted in which one group of subjects responded to the items arranged in the same random order (over 72 items for the six concepts) common to all subjects. The second group of subjects responded to the items arranged in a random order (over 72 items for the six concepts) which was a unique one for each subject of the group. No significant differences in test scores were found between the subjects receiving a common random order and those receiving a unique random order.

Tryouts of the items for item analysis and revision purposes were conducted using a single random order over the items for six concepts contained in a test booklet. This constituted a "test" of 72 items which could readily be administered in 1 hour. The tryouts were conducted during October, 1969, and January, 1970, with fifth grade students in the Madison, West Allis, and Sussex, Wisconsin school systems. Approximately 100 students (fewer for the Madison sample) responded to each "test." Madison students were given the items for six of the concepts in October; West Allis and Sussex students responded to the items for 12 concepts each in January.

The tryout data were subjected to the Generalized Item Analysis Program (GITAP) (Baker, 1969), the output of which provides the proportion responding, item-criterion biserial correlation, X_{50} (point on the criterion scale corresponding to the median of the item characteristic curve), and β (the reciprocal of the standard deviation of the item characteristic curve which is a measure of the discriminating power of the item) for each possible choice for each item as well as summary descriptive statistics for the total test. It also gives the Hoyt reliability for the total test and the standard error of measurement.

As discussed earlier, the design for these social studies achievement items is one in which the concepts and tasks are completely crossed. Since there are no item analysis procedures available for completely crossed designs, the data were analyzed in each of the two possible ways—each item as part of the appropriate concept score and as part of the appropriate task score. This raises questions as to the interpretation of such results. The main referents used for interpreting the results and as a basis for making item revisions were the results obtained from the analyses of the concept scores. The tasks were fixed and thus any arbitrary decisions were made in regard to appropriate content for incorrect choices, etc. Usual standards for item indices were not strictly adhered to, as a unique design for item analysis was being used and

a major objective of the project is to study the dimensionality of the concepts and of the tasks. If high discrimination indices were demanded, the dimensionality might have been affected by making the items more homogeneous. Also, no attempt was made to manipulate the difficulty level of the items, since another objective of the project is to determine if any differential levels of difficulty, or complexity, exist in the concepts and in the tasks. Therefore, the item analysis results were used as a very general guide to help in determining whether there were "hidden" weaknesses, clues, and/or incongruities in the items and, in an even more general sense, to show that what we were attempting to do was possible—sufficiently reliable concept and task scores could be obtained when using this completely crossed design.

The revised items can be found in "Items to Test Level of Attainment of Social Studies Concepts by Intermediate Grade Children" (Tabachnick, Weible, & Livermore, 1970).

Subjects

Pilot studies revealed that the concepts selected were very difficult for fourth graders. Thus, the decision was made to test fifth grade students with the concepts identified from the fourth grade textbooks. The social studies items were administered to 196 girls who had just completed the fifth grade during early summer, 1970, and to 195 boys who were just beginning the sixth grade during the fall of 1970 in the public school system of Madison, Wisconsin. The students were randomly selected from the population of all such girls and from the population of all such boys. The Madison Public School System made available the information concerning the populations and used their computing facilities to designate the random sample for the girls.

Initially, a random sample of 300 girls was drawn. Letters were sent to the parents of these students explaining the purpose and details of the testing, and inviting their daughter to participate in the testing program. A stamped and addressed postcard was enclosed which the parents were asked to complete and return indicating whether or not they were willing to allow their daughter to participate. One hundred and two *yes* responses and 25 *no* responses were obtained from the cards returned. Those parents who had not returned the card by a specified date were phoned. An additional 46 *yes* and 61 *no* responses were obtained by phone. Since this total of *yes* responses did not give as many subjects as were desired, an additional sample of 150 girls was drawn at random. From this sample, 56 *yes* and 30 *no* responses were obtained by card. Thus, of the total sample of 450 students, 203 *yes* and 116 *no* responses were received; seven students did not complete the testing, which resulted in a total of 196 girls tested. These students were paid $7.50 for participating.

A random sample of 756 boys was drawn and letters were sent. By mail, 420 *yes* and 87 *no* responses were obtained. Thirty-eight of the subjects did not complete the testing, resulting in 382 boys tested. Of this total, 195 boys completed the mathematics and social studies items; the others responded to language arts and science items. As with the girls, the boys who completed the testing program were paid $7.50.

Since the participation of all students comprising the random sample was impossible to attain, test score and IQ data were obtained for both the school population and those participating students from the files of the Madison Public School System for whom the information was available. Table 2 includes the summary statistics for the population of fifth grade students in the public school system of the city of Madison during the school year of 1969-70, and for the boys and the girls who comprised the tested samples for the social studies items. The Lorge-Thorndike Intelligence scores were obtained in the fall of 1968 when the subjects were fourth graders, and the scores on the Iowa Tests of Basic Skills, given in grade equivalent scores, were obtained in the fall of 1969 when the subjects were fifth graders.

Data were collected from the students using the Master Occupational Code of the United States Bureau of the Census. These data were tabulated and are presented in Table 3.

Data Collection

The data for the girls were collected in two centrally located schools, one on the East side and one on the West side of the city, during five 2-hour daily sessions for one week. Subjects could choose the week and the school in which they wanted to report for testing. A one-week session was held at Hawthorne School from June 22 to June 26, and a one-week session was held at Hoyt School from July 13 to July 17. Each 2-hour session consisted of a 72-item "test" com-

Table 2. Test Data for Population and Samples

Test		Population	Boys	Girls
Lorge-Thorndike Intelligence	$\bar{X}$	106.60	105.95	112.02
	s		14.74	12.15
	N	2605	169	191
Iowa Basic Skills				
Vocabulary	$\bar{X}$	5.53	5.60	5.75
	s		1.39	1.34
	N	2520	181	187
Reading Comprehension	$\bar{X}$	5.44	5.43	5.84
	s		1.60	1.46
	N	2520	181	187
Language Skills	$\bar{X}$	5.24	5.07	5.74
	s		1.43	1.29
	N	2520	181	187
Work-Study Skills	$\bar{X}$	5.46	5.50	5.70
	s		1.31	1.13
	N	2520	181	187
Arithmetic Skills	$\bar{X}$	5.05	5.08	5.24
	s		1.04	.97
	N	2520	179	187
Composite	$\bar{X}$	5.35	5.34	5.65
	s		1.22	1.10
	N	2520	179	185

posed of mathematics items, a 72-item "test" composed of social studies items, and an activity break between the two of approximately 1/2 hour. The social studies and the mathematics items were given first on alternate days.

The data for the boys were collected in a similar manner from mid-October to early November. Ninety of the boys who were attending Middle School for sixth grade were tested after school for five consecutive days at Schenk (October 19-23), Sennett (October 26-30), and Orchard Ridge (November 2-6) schools; those 105 elementary school boys who completed the testing were tested on three consecutive Saturday mornings (October 10, 17, and 24) at Franklin, Longfellow, and Randall schools.

The social studies items were arranged in four 72-item and one 71-item "tests." The order of the items was assigned randomly over the potential 360 items. Two different random orders were used to collect the data: one for each school for the girls and one for each type of school for the boys.

The items were arranged in five test booklets according to the random order. The students responded to the items by marking their chosen response directly on an answer sheet. The answer sheets were read by machine and the responses punched onto data cards. The tests were given by experienced test administrators.

Treatment of the Data

The treatment of the data consisted of two main procedures: reliability estimation and item analysis. The data were analyzed separately for each sex group. Hoyt analysis of variance reliability estimates were obtained for each of the 30 concept scores and each of the 12 task scores for each group studied. Means and standard deviations for each of the scores were also computed.

Item analyses using the GITAP program (Baker, 1968) were obtained for each of the items as a part of two different scores: an appropriate concept score and an appropriate task score. This program provides proportion responding, item-criterion biserial correlation, X_{50}, and β statistics for each choice of each

Table 3. Distribution of Fathers' Occupations

Occupation	Boys	Girls
PROFESSIONAL, TECHNICAL, AND KINDRED WORKERS		
00. Accountant	2	2
01. Architect	1	1
02. Dentist	—	—
03. Engineer	5	8
04. Lawyer, Judge	4	3
05. Clergyman	—	—
06. Doctor	7	4
07. Nurse	—	—
08. Teacher, Professor	18	21
09. Other Professional	16	22
FARMER		
11. Farmer	—	—
MANAGERS, OFFICIALS, PROPRIETORS, EXCEPT FARM		
21. Owner of Business	2	—
22. Manager, Official	12	11
CLERICAL AND KINDRED WORKERS		
31. Bookkeeper	—	—
32. Receptionist	—	—
39. Other Clerical and Kindred Workers	3	5
SALES WORKERS		
49. Salesman	20	15
CRAFTSMEN, FOREMEN, AND KINDRED WORKERS (SKILLED WORKERS)		
51. Craftsman, Skilled Worker	31	17
52. Foreman	2	4
53. Armed Services—Officer	1	1
54. Armed Services—Enlisted Man	1	—
OPERATIVES AND KINDRED WORKERS (SEMI-SKILLED WORKERS)		
61. Truck Driver	10	5
62. Operative in Factory	9	8
69. Other Operative and Kindred Workers	18	23
PRIVATE HOUSEHOLD AND SERVICE WORKERS		
71. Fireman	1	3
72. Policeman	1	—
73. Other Protective Service Worker	—	1
74. Practical Nurse, Nurse's Aide	2	—
75. Private Household Workers	1	—
79. Other Service Workers	14	13
81. Non-Farm Laborer	—	—
82. Farm Laborer	—	—
91. Not presently in labor force	4	8
99. Not ascertained	13	22

item. The proportion of students who respond correctly to an item is an index of the difficulty level of that item. The greater the value of the difficulty index, the easier the item. The biserial correlation coefficient is an index of the discriminating ability of the item choice. For these analyses the criterion ability used was total concept or total task score. X_{50} is the point on the criterion scale, given in standard deviation units, corresponding to the median of the item characteristic curve. It is the point at which subjects with that score have a 50-50 chance of choosing that response. β is the reciprocal of the standard deviation of the item characteristic curve at the X_{50} point. It is an index of the discrimination power of the item.

III
Results and Discussion

The means, standard deviations, and Hoyt reliability estimates obtained for the data collected during summer and fall of 1970 using the revised items are presented, separately for boys and girls, for total concept and total task scores. Also included in this section are a presentation and discussion of the item indices obtained for the correct choice of each item using both concept and task criterion scores.

Reliability Estimates and Test Statistics

Table 4 contains the means, standard deviations, and Hoyt reliability estimates obtained for the data collected during summer and fall, 1970, using the revised items for total concept scores. Table 5 contains this information for total task scores. The data were analyzed separately for the 195 boys and the 196 girls. In general, the concept scores consist of 12 items each, and the task scores of 30 items each. Exceptions to this are given in the footnotes.

The mean scores for boys are generally lower than are the mean scores for girls. No conclusions can be drawn from this, however, as the data for the girls were collected in early summer shortly after the school year of their fifth grade had ended and the data for the boys were collected in the fall shortly after the school year of their sixth grade had begun. Thus, it cannot be determined what, if any, of this difference is due to a sex difference and what is due to a time difference and possible forgetting factor. It should also be noted that the scores for the concepts Physical Feature Map and Delta are based on one more item for boys than they are for girls; Physical Feature Map and Delta have 12 items each for boys and 11 items for girls. The scores for tasks 4 and 7 are made up of 30 items for boys but only 29 for girls. The standard deviations are generally bigger for boys than they are for girls.

The reliability estimates are comparable for boys and girls. The reliability estimates for the task scores are generally a few points higher for boys than they are for girls, and the estimates for just over half of the concepts are slightly higher for boys with the remaining ones being higher for the girls.

The reliability estimates are sufficiently high to warrant study of the dimensionality of these selected social studies concepts and the tasks when using social studies content. This is a major objective of the CAA Project and is the main purpose for developing these items to measure social studies concept attainment.

As was mentioned earlier, the subject matter specialists categorized the identified social studies concepts into three major areas: Geographic Region, Man and Society, and Map and Globe Study. This was done on a theoretical basis. The data could be, and were, analyzed by area for task scores. Instead of a single total task score consisting of the score for that task type item for each of the 30 concepts, three different task scores were obtained for each of the 12 tasks, consisting of the scores for that task type item for each of the 10 concepts within a single area. The mean, standard deviation, and Hoyt reliability estimate for each of these 36 scores, 3 areas by 12 tasks, were obtained. Table 6 contains the reliability estimates obtained for task scores by area and for the total across all 30 of the concepts. Spearman-Brown estimates for tripled test lengths (some are given at the bottom of Table 6 for comparison purposes) indicate that the area distinctions are not important ones; the reliability estimates for the total task scores are about what would be

Table 4. Means, Standard Deviations, and Reliability Estimates for Social Studies Concept Scores: Boys and Girls[a]

Concept	Mean		Standard Deviation		Hoyt Reliability	
	Boys[b]	Girls[c]	Boys	Girls	Boys	Girls
1	8.71	9.22	2.53	2.34	.72	.70
2	6.93	5.97†	2.71	2.55	.71	.68
3	8.18	9.16	2.09	1.76	.54	.48
4	8.49	8.36	2.28	2.19	.63	.61
5	8.93	9.99	2.28	1.97	.65	.65
6	8.31	9.01	2.54	2.30	.69	.70
7	6.80	6.74	2.47	2.55	.62	.66
8	7.60	7.95	2.72	2.55	.71	.70
9	7.80	8.29	2.86	2.60	.75	.70
10	7.71†	7.85†	2.55	2.39	.73	.72
11	8.45	9.43	2.59	2.21	.69	.66
12	8.41	9.42	2.54	1.88	.70	.53
13	8.32	9.97	2.62	1.95	.72	.65
14	6.64	7.46	2.72	2.72	.67	.70
15	7.61	8.85	2.73	2.54	.70	.74
16	6.93	7.98	2.91	2.41	.74	.68
17	8.19	8.97	2.68	2.36	.73	.69
18	7.36	8.40	2.68	2.14	.71	.61
19	8.33	9.70	2.71	2.33	.73	.74
20	8.29	9.14	2.47	2.29	.67	.70
21	6.82	7.04	2.39	2.27	.59	.59
22	8.07	9.21	2.57	2.39	.68	.71
23	7.21	8.02	2.92	2.90	.74	.76
24	8.82	9.73	2.13	1.62	.62	.51
25	6.49	7.20	2.12	2.09	.48	.56
26	6.20	7.16	2.43	2.23	.57	.54
27	6.77	7.52	2.44	2.44	.59	.64
28	6.63	7.30	2.37	2.39	.60	.65
29	6.28	6.55†	2.80	2.52	.72	.70
30	7.04	8.16	2.50	2.13	.65	.60

[a] Scores consist of 12 items each except those marked by † which have 11 items each.

[b] N = 195

[c] N = 196

Table 5. Means, Standard Deviations, and Reliability Estimates for Social Studies Task Scores: Boys and Girls[a]

Number	Mean		Standard Deviation		Hoyt Reliability	
	Boys[b]	Girls[c]	Boys	Girls	Boys	Girls
1	22.54	24.40	5.23	4.29	.84	.81
2	21.30	23.16	5.33	4.46	.83	.80
3	22.14	24.28	5.41	4.43	.84	.81
4	22.32	23.03†	4.76	3.82	.81	.77
5	21.30	22.54	5.63	5.15	.85	.84
6	16.55	15.15	5.90	5.35	.82	.80
7	14.24	14.63†	5.60	4.78	.80	.74
8	19.29	22.14	6.55	5.67	.88	.86
9	18.61	21.34	6.44	5.71	.87	.86
10	19.25	21.90	5.49	5.26	.82	.84
11	17.04†	18.29†	4.99	4.52	.78	.76
12	13.71	15.58	5.98	5.69	.83	.82

[a]Scores consist of 30 items each except those marked by † which have 29 items each.

[b]N = 195

[c]N = 196

Key for Tasks:

1 Given name of attribute, select example.
2 Given example of attribute, select name.
3 Given name of concept, select example.
4 Given name of concept, select nonexample.
5 Given example of concept, select name.
6 Given concept, select relevant attribute.
7 Given concept, select irrelevant attribute.
8 Given definition of concept, select name.
9 Given name of concept, select definition.
10 Given concept, select supraordinate concept.
11 Given concept, select subordinate concept.
12 Given two concepts, select relationship.

Table 6. Reliability Estimates for Task Scores by Area and Total for Girls

Task	Area: Geographic Region[a]	Area: Man and Society[a]	Area: Man and Globe Study[a]	Total[b]
1	.50	.66	.59	.81
2	.51	.56	.62	.80
3	.50	.64	.58	.81
4	.47†	.60	.53	.77‡
5	.63	.63	.68	.84
6	.63	.62	.46	.80
7	.60	.54	.36†	.74‡
8	.67	.63	.70	.86
9	.68	.69	.66	.86
10	.64	.75	.60	.84
11	.53†	.62	.40	.76‡
12	.63	.58	.56	.81

[a] Scores consist of 10 items each except those marked by † which have 9 items each.

[b] Scores consist of 30 items each except those marked by ‡ which have 29 items each.

For comparison, these are the Spearman-Brown estimates for tripled test length:

Original	Estimated
.40	.67
.50	.75
.60	.82
.65	.85
.70	.88

expected from tripling the length of the test when the single area reliability estimates are of the magnitude that were obtained. Also, preliminary factor results indicate that the area distinctions are not important ones. The factor analyses of these data will be reported in a later paper.

Item Indices

Table 7 contains the item indices obtained, separately for boys and girls, based on both concept and task criterion scores. The indices included are proportion correct (this frequently is called difficulty or P), item-criterion biserial correlation, X_{50}, and β. They are given for the correct choice only. The key for the concepts is given in Appendix A and the key for the tasks is given in Table 5. The item number has no special meaning; it is a coding number and was included in the table as an organizational aid. Note that proportion correct is the same whether analyzed using the concept criterion score or the task criterion score; hence, there is only one column each for boys and girls. The other item indices differ according to criterion score used. When an item was missing from the data collected, the appropriate row was left blank except for the identifying numbers. This is the case for Item 119 for Concept 10, Task 11. Two items, Nos. 16 and 343, were missing from the data collected for the girls but were available for the boys; in this case only the columns for the girls are blank. There are a few instances where there is a blank in an X_{50} column. If β is very low, the X_{50} becomes meaningless; thus, X_{50} is not included if the β value is less than .05. There are two items, Nos. 52 and 145, for which the X_{50} and β columns for girls' concept scores are blank. The biserial

Table 7. Item Indices Based on Concept and Task Criterion Scores

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
1	1	1	.83	.77	.60	.38	.71	.52	-1.60	-2.47	-1.04	-1.43	.74	.42	1.01	.61
	2	2	.81	.85	.56	.62	.54	.54	-1.54	-1.42	-1.93	-1.9.	.69	.79	.64	.65
	3	3	.90	.96	.64	.70	.71	.71	-2.02	-1.84	-2.55	-2.51	.83	.99	1.00	1.03
	4	4	.95	.97	.55	.67	.23	.29	-3.05	-2.53	-8.13	-6.38	.66	.89	.24	.31
Coastline	5	5	.78	.77	.82	.67	.77	.76	- .94	-1.16	- .94	- .95	1.45	.90	1.19	1.18
	6	6	.52	.57	.71	.58	.63	.51	- .08	- .10	- .29	- .35	1.01	.72	.81	.59
	7	7	.65	.69	.75	.67	.72	.60	- .52	- .58	- .70	- .84	1.12	.89	1.04	.75
	8	8	.68	.65	.67	.57	.77	.64	- .69	- .80	- .51	- .61	.90	.70	1.19	.84
	9	9	.64	.80	.81	.74	.86	.80	- .43	- .47	- .99	-1.06	1.36	1.09	1.65	1.32
	10	10	.77	.89	.66	.61	.66	.64	-1.12	-1.22	-1.84	-1.89	.87	.76	.88	.84
	11	11	.68	.72	.69	.62	.75	.70	- .67	- .74	- .77	- .82	.94	.80	1.14	.99
	12	12	.50	.57	.64	.57	.72	.65	.01	.01	- .25	- .27	.84	.69	1.03	.87
2	1	13	.52	.60	.76	.64	.51	.47	- .06	- .07	- .48	- .52	1.16	.83	.60	.53
	2	14	.69	.73	.55	.55	.58	.37	- .91	- .91	-1.06	-1.66	.67	.66	.71	-.27
	3	15	.77	.73	.58	.52	.71	.52	-1.30	-1.45	- .89	-1.22	.71	.61	1.00	.60
	4	16														
	5	17	.64	.54	.73	.65	.79	.58	- .47	- .54	- .13	- .18	1.07	.85	1.27	.72
Delta	6	18	.70	.63	.66	.57	.67	.55	- .81	- .93	- .51	- .61	.87	.69	.90	.66
	7	19	.42	.44	.59	.57	.57	.44	.34	.35	.25	.32	.74	.70	.70	.49
	8	20	.57	.64	.72	.68	.69	.80	- .26	- .28	- .53	- .46	1.03	.93	.96	1.31
	9	21	.61	.59	.66	.37	.59	.59	- .40	- .72	- .40	- .39	.89	.40	.73	.73
	10	22	.57	.42	.62	.39	.77	.53	- .30	- .48	.27	.39	.79	.43	1.20	.62
	11	23	.14	.18	.38	.30	.20	.13	2.90	3.66	4.62	7.15	.40	.31	.20	.13
	12	24	.48	.46	.66	.60	.72	.66	.09	.10	.14	.16	.88	.75	1.03	.88
3	1	25	.80	.88	.66	.55	.56	.49	-1.27	-1.52	-2.09	-2.37	.88	.66	.67	.56
	2	26	.70	.73	.61	.65	.67	.66	- .85	- .79	- .91	- .93	.76	.86	.91	.87
	3	27	.80	.90	.63	.44	.41	.22	-1.35	-1.93	-3.08	-5.67	.80	.48	.45	.23
	4	28	.96	.97	.41	.51	.68	.61	-4.34	-3.55	-2.89	-3.20	.46	.59	.92	.77
	5	29	.91	.97	.56	.44	.48	.28	-2.42	-3.10	-4.07	-7.02	.68	.49	.55	.29
Desert	6	30	.48	.64	.53	.48	.60	.34	.09	.09	- .61	-1.06	.62	.55	.76	.37
	7	31	.41	.40	.58	.53	.74	.50	.41	.46	.35	.52	.71	.62	1.10	.57
	8	32	.75	.91	.55	.66	.73	.58	-1.26	-1.05	-1.86	-2.36	.65	.87	1.07	.71
	9	33	.67	.84	.77	.74	.72	.67	- .58	- .60	-1.36	-1.47	1.20	1.12	1.04	.90
	10	34	.77	.86	.60	.48	.60	.66	-1.26	-1.57	-1.82	-1.65	.75	.55	.75	.88
	11	35	.58	.67	.48	.38	.48	.34	- .42	- .53	- .94	-1.32	.55	.41	.55	.36
	12	36	.34	.38	.35	.26	.42	.16	1.14	1.58	.72	1.83	.38	.26	.46	.17
4	1	37	.93	.94	.70	.66	.64	.67	-2.08	-2.21	-2.49	-2.38	.99	.88	.83	.90
	2	38	.89	.94	.56	.57	.45	.40	-2.20	-2.18	-3.44	-3.90	.68	.69	.50	.43
	3	39	.91	.84	.73	.75	.73	.61	-1.81	-1.77	-1.31	-1.63	1.07	1.13	1.18	.78
	4	40	.89	.93	.57	.62	.65	.65	-2.11	-1.96	-2.33	-2.32	.70	.79	.85	.85
	5	41	.56	.51	.56	.47	.49	.42	- .29	- .34	- .03	- .03	.67	.53	.58	.47
Gulf	6	42	.60	.67	.61	.52	.63	.52	- .41	- .49	- .72	- .86	.78	.61	.80	.62
	7	43	.58	.64	.54	.54	.67	.70	- .37	- .37	- .52	- .50	.65	.64	.91	.99
	8	44	.52	.54	.62	.56	.53	.46	- .07	- .08	- .19	- .22	.78	.67	.62	.51
	9	45	.64	.56	.60	.51	.68	.55	- .60	- .70	- .23	- .28	.75	.60	.92	.66
	10	46	.74	.86	.67	.54	.78	.61	- .98	-1.22	-1.40	-1.78	.90	.64	1.25	.78
	11	47	.72	.37	.73	.69	.52	.42	- .79	- .83	.65	.80	1.08	.96	.61	.47
	12	48	.51	.55	.59	.54	.66	.49	- .05	- .06	- .19	- .26	.73	.65	.89	.56

(continued)

Table 7 (continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
5 Mountain Region	1	49	.78	.84	.41	.38	.56	.50	-1.94	-2.05	-1.74	-1.98	.45	.42	.68	.57
	2	50	.78	.87	.63	.67	.84	.67	-1.23	-1.16	-1.32	-1.66	.81	.89	1.56	.91
	3	51	.86	.92	.60	.56	.59	.62	-1.79	-1.91	-2.41	-2.31	.74	.67	.73	.79
	4	52	.97	.98	.70	.57	1.21	.79	-2.78	-3.42		-2.59	.98	.69		1.29
	5	53	.90	.94	.56	.47	.68	.70	-2.33	-2.73	-2.26	-2.21	.67	.54	.93	.98
	6	54	.48	.63	.69	.63	.66	.65	.08	.09	- .51	- .52	.94	.81	.88	.86
	7	55	.64	.73	.64	.59	.61	.60	- .57	- .61	-1.00	-1.02	.83	.74	.78	.75
	8	56	.71	.86	.69	.58	.83	.80	- .80	- .95	-1.31	-1.36	.94	.70	1.50	1.33
	9	57	.70	.84	.71	.62	.55	.57	- .75	- .86	-1.77	-1.72	1.02	.79	.66	.70
	10	58	.77	.89	.73	.64	.85	.84	-1.03	-1.18	-1.42	-1.45	1.07	.83	1.63	1.53
	11	59	.73	.67	.67	.52	.71	.62	- .94	-1.21	- .61	- .70	.89	.60	1.01	.79
	12	60	.59	.83	.67	.53	.86	.71	- .36	- .45	-1.10	-1.32	.91	.63	1.65	1.01
6 River	1	61	.67	.79	.73	.77	.72	.70	- .59	- .56	-1.10	-1.14	1.07	1.21	1.04	.97
	2	62	.75	.86	.55	.42	.78	.70	-1.23	-1.62	-1.39	-1.55	.67	.47	1.26	.99
	3	63	.83	.88	.58	.63	.58	.58	-1.65	-1.52	-2.05	-2.06	.71	.81	.71	.70
	4	64	.84	.95	.50	.41	.47	.39	-1.96	-2.36	-3.71	-4.52	.58	.45	.53	.42
	5	65	.84	.87	.64	.55	.78	.54	-1.56	-1.80	-1.42	-2.05	.83	.67	1.26	.65
	6	66	.43	.38	.50	.48	.60	.54	.35	.37	.52	.58	.57	.54	.75	.63
	7	67	.55	.46	.68	.52	.53	.36	- .20	- .26	.17	.25	.93	.61	.63	.39
	8	68	.75	.71	.57	.47	.74	.64	-1.16	-1.43	- .77	- .88	.71	.53	1.08	.83
	9	69	.67	.83	.76	.68	.73	.70	- .56	- .64	-1.31	-1.37	1.18	.92	1.08	.99
	10	70	.69	.86	.63	.57	.83	.84	- .80	- .88	-1.31	-1.30	.80	.70	1.52	1.54
	11	71	.78	.78	.74	.67	.88	.76	-1.00	-1.15	- .88	-1.02	1.10	.91	1.83	1.17
	12	72	.51	.67	.73	.72	.72	.60	- .04	- .04	- .62	- .75	1.06	1.03	1.04	.75
7 River Mouth	1	73	.48	.42	.61	.39	.64	.43	.10	.15	.32	.48	.76	.43	.83	.47
	2	74	.28	.21	.40	.22	.27	.32	1.44	2.65	2.98	2.45	.44	.22	.28	.34
	3	75	.87	.92	.48	.66	.55	.70	-2.34	-1.72	-2.62	-2.05	.55	.88	.65	.98
	4	76	.84	.81	.51	.55	.45	.64	-1.95	-1.82	-1.90	-1.34	.59	.66	.51	.84
	5	77	.65	.59	.56	.43	.59	.42	- .69	- .91	- .40	- .55	.68	.47	.72	.47
	6	78	.57	.64	.62	.59	.76	.54	- .28	- .29	- .48	- .68	.79	.73	1.16	.63
	7	79	.44	.47	.50	.41	.53	.45	.30	.36	.14	.17	.57	.45	.63	.50
	8	80	.62	.55	.65	.35	.78	.53	- .45	- .85	- .15	- .22	.85	.37	1.24	.62
	9	81	.62	.63	.77	.59	.79	.47	- .38	- .50	- .41	- .69	1.23	.73	1.31	.55
	10	82	.43	.53	.34	.33	.44	.34	.54	.57	- .15	- .19	.37	.35	.49	.36
	11	83	.52	.43	.66	.54	.62	.49	- .09	- .11	.29	.37	.88	.64	.79	.56
	12	84	.48	.55	.66	.53	.66	.48	.07	.08	- .19	- .27	.89	.63	.88	.55
8 Strait	1	85	.78	.84	.52	.54	.60	.56	-1.52	-1.46	-1.63	-1.70	.60	.64	.76	.71
	2	86	.78	.76	.64	.58	.51	.58	-1.20	-1.32	-1.48	-1.30	.84	.72	.60	.71
	3	87	.84	.87	.67	.69	.65	.39	-1.50	-1.45	-1.73	-2.95	.89	.94	.88	.42
	4	88	.80	.89	.68	.57	.47	.40	-1.23	-1.49	-2.58	-3.01	.94	.69	.53	.44
	5	89	.65	.57	.75	.71	.77	.62	- .50	- .53	- .22	- .27	1.14	1.00	1.22	.80
	6	90	.59	.59	.43	.35	.52	.54	- .56	- .69	- .42	- .40	.48	.37	.61	.65
	7	91	.45	.49	.64	.60	.54	.47	.21	.23	.05	.05	.84	.75	.65	.53
	8	92	.67	.67	.81	.57	.74	.42	- .53	- .75	- .61	-1.06	1.40	.70	1.09	.47
	9	93	.57	.68	.68	.51	.83	.74	- .26	- .34	- .56	- .63	.94	.59	1.46	1.10
	10	94	.62	.78	.68	.53	.65	.54	- .45	- .58	-1.16	-1.40	.92	.62	.86	.64
	11	95	.41	.38	.61	.59	.63	.54	.40	.41	.48	.55	.76	.72	.80	.65
	12	96	.45	.43	.64	.58	.75	.66	.21	.23	.24	.27	.84	.72	1.12	.87

(continued)

Table 7 (continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
9 Tributary	1	97	.87	.86	.40	.51	.53	.46	-2.83	-2.24	-2.00	-2.32	.44	.59	.63	.52
	2	98	.52	.66	.33	.26	.40	.38	- .17	- .23	-1.06	-1.10	.35	.27	.44	.42
	3	99	.75	.77	.68	.59	.66	.54	- .98	-1.13	-1.10	-1.35	.93	.74	.87	.63
	4	100	.64	.73	.68	.51	.69	.63	- .51	- .69	- .89	- .97	.92	.59	.94	.81
	5	101	.81	.87	.74	.75	.77	.67	-1.16	-1.15	-1.44	-1.67	1.11	1.12	1.22	.90
	6	102	.59	.59	.75	.64	.72	.60	- .32	- .37	- .30	- .37	1.15	.84	1.05	.74
	7	103	.53	.58	.66	.63	.55	.57	- .11	- .11	- .38	- .36	.88	.80	.65	.69
	8	104	.67	.67	.74	.68	.70	.50	- .60	- .65	- .62	- .87	1.11	.93	.99	.58
	9	105	.65	.68	.81	.73	.79	.60	- .46	- .51	- .59	- .77	1.41	1.06	1.30	.75
	10	106	.64	.70	.77	.65	.66	.59	- .47	- .55	- .81	- .91	1.19	.86	.89	.74
	11	107	.65	.71	.79	.67	.77	.62	- .47	- .56	- .72	- .88	1.29	.90	1.19	.80
	12	108	.49	.48	.62	.63	.49	.34	.05	.05	.10	.15	.80	.81	.56	.36
10 Tropical Region	1	109	.84	.85	.77	.73	.52	.45	-1.30	-1.37	-2.00	-2.34	1.21	1.07	.61	.50
	2	110	.83	.94	.69	.70	.40	.60	-1.39	-1.36	-3.94	-2.63	.95	.99	.44	.76
	3	111	.76	.81	.76	.68	.87	.79	- .93	-1.03	-1.01	-1.12	1.16	.94	1.78	1.29
	4	112	.82	.82	.81	.69	.77	.66	-1.11	-1.30	-1.17	-1.37	1.38	.96	1.20	.88
	5	113	.71	.75	.84	.79	.83	.79	- .65	- .69	- .91	- .96	1.54	1.28	1.49	1.30
	6	114	.61	.61	.63	.61	.76	.59	- .42	- .44	- .37	- .48	.81	.76	1.17	.74
	7	115	.49	.37	.44	.35	.53	.39	.04	.06	.64	.88	.49	.37	.63	.42
	8	116	.72	.79	.65	.68	.80	.79	- .91	- .87	-1.01	-1.02	.85	.93	1.34	1.29
	9	117	.58	.48	.77	.66	.68	.52	- .26	- .31	.06	.07	1.21	.87	.92	.60
	10	118	.76	.74	.83	.72	.77	.75	- .84	- .99	- .86	- .88	1.51	1.03	1.20	1.12
	11	119														
	12	120	.59	.65	.63	.56	.76	.61	- .38	- .43	- .50	- .62	.80	.68	1.16	.77
11 Airway	1	121	.84	.96	.79	.65	.86	.99	-1.26	-1.53	-2.03	-1.75	1.29	.86	1.66	.75
	2	122	.71	.85	.55	.50	.74	.68	-1.02	-1.13	-1.38	-1.51	.66	.57	1.11	.92
	3	123	.62	.68	.53	.53	.65	.66	- .58	- .58	- .71	- .70	.62	.63	.86	.87
	4	124	.66	.78	.64	.56	.61	.58	- .63	- .72	-1.24	-1.31	.83	.67	.77	.71
	5	125	.73	.67	.69	.64	.71	.67	- .91	- .98	- .61	- .65	.95	.83	1.02	.89
	6	126	.66	.84	.62	.56	.57	.36	- .65	- .72	-1.76	-2.76	.78	.68	.69	.39
	7	127	.64	.64	.59	.58	.52	.42	- .61	- .62	- .68	- .85	.73	.72	.60	.46
	8	128	.80	.88	.71	.67	.80	.82	-1.19	-1.26	-1.45	-1.42	1.00	.90	1.34	1.42
	9	129	.73	.79	.73	.62	.79	.63	- .83	- .98	-1.02	-1.29	1.07	.79	1.31	.80
	10	130	.73	.82	.67	.56	.79	.69	- .92	-1.11	-1.13	-1.30	.91	.68	1.32	.97
	11	131	.75	.89	.53	.37	.55	.58	-1.27	-1.79	-2.25	-2.14	.62	.40	.66	.71
	12	132	.57	.65	.67	.55	.63	.50	- .28	- .34	- .60	- .76	.90	.66	.81	.57
12 City	1	133	.93	.98	.65	.68	.35	.18	-2.27	-2.14	-5.56	-11.13	.84	.94	.40	.19
	2	134	.87	.95	.68	.60	.75	.67	-1.63	-1.86	-2.24	-2.52	.93	.75	1.15	.90
	3	135	.74	.92	.84	.83	.70	.51	- .78	- .79	-2.00	-2.71	1.53	1.49	.97	.60
	4	136	.79	.87	.69	.53	.77	.68	-1.16	-1.51	-1.47	-1.66	.96	.63	1.22	.94
	5	137	.83	.85	.60	.69	.74	.80	-1.55	-1.36	-1.41	-1.31	.76	.95	1.10	1.33
	6	138	.61	.69	.55	.51	.61	.55	- .49	- .53	- .81	- .90	.65	.59	.77	.65
	7	139	.58	.62	.55	.41	.41	.41	- .36	- .49	- .77	- .77	.67	.45	.45	.45
	8	140	.61	.73	.68	.66	.67	.55	- .41	- .43	- .91	-1.11	.93	.87	.91	.66
	9	141	.55	.64	.68	.66	.61	.56	- .20	- .21	- .58	- .63	.93	.87	.76	.68
	10	142	.66	.78	.74	.56	.72	.66	- .56	- .75	-1.05	-1.15	1.10	.67	1.04	.88
	11	143	.74	.82	.49	.38	.47	.41	-1.33	-1.74	-1.94	-2.24	.56	.41	.54	.45
	12	144	.50	.57	.69	.63	.47	.28	.01	.01	- .36	- .60	.96	.81	.53	.29

(continued)

Table 7 (continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
Countryside	1	145	.72	.87	.82	.75	1.01	.84	- .72	- .79		-1.35	1.44	1.14		1.57
	2	146	.79	.94	.65	.65	.67	.62	-1.26	-1.27	-2.39	-2.54	.86	.85	.89	.80
	3	147	.89	.96	.68	.68	.62	.50	-1.81	-1.82	-2.80	-3.46	.94	.93	.79	.58
	4	148	.87	.93	.62	.62	.89	.62	-1.79	-1.80	-1.70	-2.44	.79	.78	1.93	.78
	5	149	.89	.94	.62	.49	.73	.61	-1.99	-2.53	-2.16	-2.59	.79	.56	1.08	.77
	6	150	.62	.74	.69	.54	.71	.60	- .42	- .54	- .93	-1.09	.96	.64	1.00	.76
	7	151	.56	.71	.79	.65	.62	.64	- .19	- .23	- .91	- .89	1.30	.85	.79	.83
	8	152	.64	.86	.55	.49	.31	.12	- .64	- .71	-3.49	-9.25	.65	.56	.33	.12
	9	153	.56	.66	.61	.57	.61	.56	- .26	- .28	- .69	- .75	.78	.69	.77	.67
	10	154	.42	.63	.68	.57	.77	.62	.29	.35	- .44	- .54	.94	.70	1.19	.80
	11	155	.63	.86	.71	.54	.73	.40	- .47	- .62	-1.50	-2.74	1.00	.64	1.07	.43
	12	156	.73	.84	.64	.49	.92	.66	- .94	-1.23	-1.08	-1.52	.84	.57	2.41	.88
Democracy	1	157	.56	.73	.65	.60	.66	.71	- .25	- .27	- .92	- .86	.86	.76	.89	1.02
	2	158	.58	.67	.69	.62	.76	.62	- .29	- .32	- .57	- .71	.94	.78	1.18	.78
	3	159	.62	.69	.62	.42	.74	.66	- .49	- .74	- .66	- .75	.80	.46	1.13	.87
	4	160	.52	.64	.49	.45	.54	.47	- .12	- .13	- .66	- .75	.56	.50	.64	.53
	5	161	.73	.73	.60	.41	.65	.56	-1.01	-1.47	- .96	-1.13	.75	.45	.86	.67
	6	162	.50	.56	.49	.33	.52	.40	.01	.02	- .27	- .35	.56	.35	.61	.43
	7	163	.39	.39	.47	.24	.21	.11	.60	1.15	1.35	2.56	.53	.25	.22	.11
	8	164	.56	.64	.71	.64	.71	.55	- .23	- .25	- .51	- .67	1.02	.83	1.02	.66
	9	165	.55	.69	.65	.62	.74	.63	- .19	- .20	- .66	- .79	.86	.80	1.11	.80
	10	166	.77	.84	.57	.57	.85	.76	-1.32	-1.32	-1.18	-1.31	.70	.69	1.60	1.18
	11	167	.43	.45	.68	.56	.63	.59	.26	.31	.18	.20	.92	.68	.82	.73
	12	168	.42	.43	.51	.37	.66	.57	.42	.58	.27	.32	.59	.40	.88	.69
Exchange	1	169	.58	.79	.71	.64	.67	.68	- .28	- .31	-1.17	-1.16	1.02	.83	.91	.93
	2	170	.64	.64	.57	.51	.68	.71	- .63	- .71	- .54	- .51	.69	.59	.93	1.02
	3	171	.66	.87	.59	.59	.80	.67	- .69	- .68	-1.42	-1.70	.72	.73	1.35	.90
	4	172	.63	.72	.66	.58	.67	.67	- .48	- .56	- .88	- .89	.88	.71	.91	.91
	5	173	.56	.64	.69	.53	.72	.57	- .24	- .30	- .49	- .62	.94	.63	1.03	.69
	6	174	.58	.67	.51	.44	.64	.60	- .39	- .45	- .70	- .75	.59	.49	.83	.75
	7	175	.56	.62	.43	.50	.74	.62	- .37	- .33	- .40	- .48	.48	.57	1.11	.80
	8	176	.74	.89	.83	.82	.77	.72	- .77	- .78	-1.58	-1.69	1.46	1.42	1.20	1.03
	9	177	.71	.88	.63	.65	.93	.87	- .89	- .87	-1.28	-1.37	.81	.85	2.44	1.73
	10	178	.70	.81	.76	.62	.88	.78	- .68	- .84	-1.00	-1.13	1.16	.78	1.87	1.26
	11	179	.84	.90	.73	.65	.68	.55	-1.36	-1.54	-1.87	-2.31	1.08	.86	.92	.6
	12	180	.41	.42	.51	.44	.54	.48	.45	.52	.38	.43	.59	.49	.64	.54
Government	1	181	.72	.72	.71	.73	.66	.59	- .84	- .81	- .90	-1.02	1.00	1.07	.88	.72
	2	182	.67	.74	.58	.56	.51	.35	- .75	- .78	-1.30	-1.87	.71	.67	.59	.38
	3	183	.49	.27	.73	.63	.76	.66	.03	.03	- .24	- .27	1.08	.80	1.18	.88
	4	184	.34	.28	.45	.42	.48	.45	.90	.96	1.21	1.29	.50	.46	.55	.50
	5	185	.61	.76	.56	.56	.60	.52	- .48	- .47	-1.18	-1.37	.67	.68	.75	.60
	6	186	.57	.77	.68	.57	.68	.63	- .26	- .31	-1.08	-1.18	.92	.69	.94	.81
	7	187	.29	.24	.62	.42	.37	.31	.91	1.33	1.90	2.31	.78	.46	.40	.32
	8	188	.55	.68	.75	.63	.77	.71	- .18	- .22	- .62	- .67	1.14	.80	1.22	1.01
	9	189	.56	.71	.79	.72	.73	.71	- .21	- .22	- .78	- .79	1.27	1.04	1.06	1.02
	10	190	.79	.90	.69	.65	.85	.84	-1.20	-1.26	-1.54	-1.55	.95	.87	1.59	1.53
	11	191	.74	.90	.66	.61	.72	.69	- .99	-1.08	-1.76	-1.85	.87	.76	1.04	.95
	12	192	.58	.69	.74	.66	.72	.49	- .29	- .33	- .71	-1.03	1.11	.87	1.03	.57

(continued)

GPO 827-034-4

Table 7 (continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
17 Land Routes	1	193	.72	.79	.57	.49	.69	.66	-1.03	-1.20	-1.15	-1.19	.70	.56	.94	.89
	2	194	.77	.83	.69	.68	.58	.58	-1.07	-1.08	-1.65	-1.65	.95	.93	.72	.72
	3	195	.74	.81	.59	.63	.80	.67	-1.09	-1.01	-1.08	-1.30	.72	.82	1.34	.89
	4	196	.82	.88	.72	.60	.78	.73	-1.27	-1.53	-1.49	-1.59	1.04	.75	1.25	1.07
	5	197	.42	.41	.48	.29	.42	.33	.44	.74	.52	.67	.55	.30	.46	.35
	6	198	.59	.72	.85	.69	.72	.54	- .28	- .35	- .81	-1.08	1.64	.97	1.04	.63
	7	199	.53	.59	.65	.55	.65	.61	- .13	- .15	- .36	- .39	.85	.65	.86	.78
	8	200	.75	.85	.77	.76	.66	.72	- .87	- .88	-1.57	-1.45	1.21	1.17	.89	1.04
	9	201	.73	.74	.66	.55	.64	.[illegible]	- .95	-1.13	-1.00	-1.22	.87	.66	.84	.62
	10	202	.81	.88	.81	.58	.78	.74	-1.07	-1.49	-1.52	-1.61	1.36	.71	1.26	1.09
	11	203	.81	.80	.66	.58	.92	.76	-1.33	-1.51	- .92	-1.11	.89	.72	2.34	1.18
	12	204	.50	.67	.63	.51	.53	.38	.01	.01	- .82	-1.13	.80	.59	.62	.42
18 News	1	205	.48	.61	.55	.38	.65	.61	.11	.15	- .42	- .44	.65	.41	.86	.77
	2	206	.83	.93	.72	.77	.36	.46	-1.30	-1.21	-4.06	-3.16	1.05	1.21	.39	.52
	3	207	.56	.70	.66	.57	.53	.46	- .25	- .28	- .98	-1.14	.87	.69	.63	.52
	4	208	.25	.17	.40	.31	.63	.50	1.72	2.21	1.52	1.92	.44	.33	.81	.58
	5	209	.72	.89	.66	.68	.70	.62	- .88	- .85	-1.74	-1.97	.87	.92	.97	.78
	6	210	.64	.69	.75	.66	.71	.54	- .48	- .55	- .71	- .93	1.14	.87	1.02	.65
	7	211	.47	.49	.48	.50	.46	.35	.18	.17	.06	.07	.54	.58	.52	.38
	8	212	.69	.81	.80	.72	.77	.79	- .63	- .70	-1.13	-1.09	1.32	1.03	1.19	1.30
	9	213	.81	.86	.64	.53	.81	.65	-1.37	-1.65	-1.33	-1.64	.84	.63	1.36	.86
	10	214	.77	.90	.72	.62	.68	.71	-1.04	-1.22	-1.92	-1.82	1.05	.79	.92	1.02
	11	215	.70	.77	.76	.57	.64	.62	- .68	- .90	-1.15	-1.20	1.16	.70	.84	.78
	12	216	.45	.59	.63	.59	.56	.58	.21	.23	- .42	- .40	.82	.73	.68	.71
19 Organization	1	217	.75	.91	.61	.47	.96	.62	-1.12	-1.46	-1.42	-2.21	.77	.53	3.29	.78
	2	218	.74	.72	.61	.51	.68	.54	-1.08	-1.28	- .85	-1.08	.77	.60	.93	.64
	3	219	.71	.83	.66	.60	.79	.65	- .85	- .94	-1.18	-1.45	.89	[illegible]	1.31	.85
	4	220	.71	.78	.70	.52	.87	.61	- .80	-1.07	- .87	-1.24	.99	[illegible]	1.77	.77
	5	221	.73	.85	.70	.61	.76	.77	- .87	- .99	-1.37	-1.35	.98	.78	1.18	1.22
	6	222	.62	.78	.80	.68	.87	.77	- .38	- .45	- .89	-1.01	1.34	.93	1.76	1.21
	7	223	.55	.76	.68	.61	.72	.50	- .18	- .20	- .95	-1.37	.92	.78	1.05	.88
	8	224	.79	.93	.65	.58	1.01	.74	-1.25	-1.40	.00	-2.03	.85	.70	.00	1.10
	9	225	.75	.89	.80	.77	.66	.52	- .84	- .87	-1.84	-2.33	1.32	1.21	.88	.61
	10	226	.72	.84	.62	.56	.72	.66	- .95	-1.06	-1.37	-1.50	.79	.68	1.03	.87
	11	227	.80	.85	.80	.70	.69	.68	-1.05	-1.21	-1.52	-1.53	1.35	.97	.95	.94
	12	228	.45	.57	.41	.36	.56	.52	.33	.36	- .32	- .35	.45	.38	.67	.61
20 Waterway	1	229	.86	.91	.60	.58	.73	.71	-1.81	-1.86	-1.87	-1.93	.75	.72	1.06	1.00
	2	230	.81	.88	.72	.66	.90	.58	-1.20	-1.31	-1.30	-2.01	1.03	.87	2.01	.71
	3	231	.67	.82	.56	.41	.84	.74	- .77	-1.05	-1.07	-1.22	.68	.45	1.55	1.11
	4	232	.72	.82	.69	.62	.58	.67	- .84	- .93	-1.55	-1.35	.95	.79	.71	.90
	5	233	.80	.88	.55	.57	.77	.71	-1.52	-1.48	-1.50	-1.65	.66	.69	1.22	1.00
	6	234	.59	.65	.77	.62	.58	.42	- .29	- .37	- .67	- .94	1.21	.78	.72	.46
	7	235	.52	.51	.69	.61	.69	.66	- .08	- .10	- .04	- .04	.95	.76	.94	.88
	8	236	.66	.78	.62	.62	.72	.58	- .65	- .65	-1.06	-1.31	.79	.78	1.03	.71
	9	237	.82	.92	.67	.62	.74	.74	-1.37	-1.48	-1.88	-1.88	.90	.79	1.11	1.11
	10	238	.70	.78	.69	.57	.78	.81	- .77	- .94	.97	- .93	.95	.69	1.26	1.40
	11	239	.78	.84	.59	.63	.64	.59	-1.31	-1.22	-1.56	-1.71	.72	.81	.84	.73
	12	240	.37	.37	.37	.37	.63	.59	.89	.89	.54	.57	.40	.40	.82	.74

(continued)

Table 7 (continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
21	1	241	.79	.87	.63	.76	.80	.81	-1.30	-1.09	-1.43	-1.40	.82	1.16	1.32	1.40
	2	242	.85	.89	.56	.55	.71	.58	-1.83	-1.86	-1.72	-2.08	.67	.66	1.00	.72
	3	243	.67	.51	.61	.36	.65	.51	- .70	-1.21	- .02	- .03	.78	.38	.85	.59
	4	244	.53	.49	.83	.77	.67	.60	- .10	- .11	.04	.04	1.51	1.22	.90	.75
	5	245	.73	.72	.71	.57	.85	.73	- .56	-1.06	- .70	- .82	1.00	.70	1.61	1.07
Country	6	246	.49	.66	.45	.42	.61	.56	.04	.05	- .69	- .76	.50	.47	.76	.67
	7	247	.33	.26	.34	.16	.37	.12	1.29	2.71	1.75	5.47	.37	.17	.39	.12
	8	248	.50	.54	.62	.55	.67	.60	- .01	- .01	- .13	- .15	.79	.66	.89	.75
	9	249	.58	.72	.67	.53	.66	.72	- .32	- .40	- .90	- .83	.90	.62	.88	1.04
	10	250	.68	.78	.59	.56	.37	.43	- .81	- .85	-2.03	-1.78	.73	.68	.40	.47
	11	251	.32	.24	.19	.03	.11	-.05	2.43		6.02		.19	.03	.12	-.05
	12	252	.33	.36	.51	.42	.47	.30	.84	1.02	.78	1.23	.60	.47	.53	.31
22	1	253	.82	.88	.73	.67	.77	.66	-1.23	-1.35	-1.55	-1.80	1.07	.87	1.19	.88
	2	254	.76	.89	.72	.62	.71	.68	-1.00	-1.17	-1.71	-1.78	1.04	.78	1.01	.93
	3	255	.70	.88	.63	.62	.93	.93	- .84	- .85	-1.24	-1.25	.82	.80	2.63	2.52
	4	256	.75	.88	.74	.56	.61	.57	- .92	-1.23	-1.93	-2.09	1.11	.67	.78	.69
	5	257	.76	.82	.71	.73	.72	.71	- .99	- .96	-1.26	-1.26	1.02	1.07	1.03	1.02
	6	258	.75	.72	.56	.46	.75	.62	-1.20	-1.46	- .79	- .97	.68	.52	1.15	.78
Distance	7	259	.44	.55	.54	.57	.67	.63	.30	.28	- .17	- .18	.65	.69	.90	.82
	8	260	.76	.84	.60	.56	.72	.60	-1.18	-1.26	-1.39	-1.68	.74	.67	1.03	.75
	9	261	.69	.78	.57	.47	.63	.63	- .88	-1.07	-1.21	-1.19	.70	.53	.81	.82
	10	262	.64	.76	.61	.60	.73	.61	- .59	- .60	- .97	-1.17	.77	.75	1.05	.76
	11	263	.58	.68	.55	.47	.60	.36	- .39	- .46	- .80	-1.33	.66	.53	.75	.39
	12	264	.41	.53	.61	.63	.70	.63	.37	.36	- .11	- .12	.78	.81	.97	.82
23	1	265	.80	.77	.68	.64	.63	.60	-1.24	-1.31	-1.17	-1.23	.92	.84	.82	.75
	2	266	.74	.73	.69	.61	.67	.59	- .92	-1.05	- .93	-1.06	.96	.77	.91	.73
	3	267	.72	.87	.73	.72	.79	.72	- .79	- .80	-1.41	-1.56	1.08	1.03	1.30	1.02
	4	268	.65	.70	.63	.61	.75	.58	- .52	- .62	- .72	- .92	1.06	.76	1.13	.72
	5	269	.50	.48	.59	.50	.70	.59	.01	.01	.05	.06	.72	.58	.99	.73
	6	270	.57	.60	.72	.64	.69	.59	- .24	- .27	- .37	- .44	1.05	.83	.96	.73
Latitude	7	271	.35	.45	.61	.65	.54	.44	.61	.58	.24	.29	.78	.85	.65	.49
	8	272	.65	.74	.65	.62	.77	.58	- .60	- .63	- .83	-1.11	.85	.78	1.21	.71
	9	273	.54	.63	.41	.28	.68	.58	- .24	- .35	- .48	- .56	.45	.29	.92	.72
	10	274	.53	.63	.68	.58	.68	.65	- .10	- .12	- .51	- .52	.93	.73	.92	.85
	11	275	.56	.71	.66	.62	.79	.62	- .25	- .26	- .72	- .91	.87	.80	1.28	.80
	12	276	.60	.69	.76	.67	.72	.60	- .33	- .38	- .71	- .84	1.20	.91	1.03	.75
24	1	277	.89	.86	.58	.56	.73	.56	-2.10	-2.16	-1.46	-1.91	.71	.68	1.07	.68
	2	278	.75	.87	.77	.68	.79	.75	- .89	-1.01	-1.45	-1.53	1.21	.93	1.28	1.12
	3	279	.88	.95	.59	.32	.54	.83	-2.00	-3.67	-3.12	-2.03	.74	.34	.64	1.49
	4	280	.71	.95	.72	.77	.40	.48	-1.84	-1.72	-4.10	-3.44	1.04	1.22	.44	.54
	5	281	.79	.86	.70	.72	.72	.52	-1.16	-1.13	-1.50	-2.08	.97	1.02	1.05	.61
	6	282	.76	.84	.59	.46	.74	.54	-1.19	-1.53	-1.36	-1.86	.73	.52	1.09	.64
Globe	7	283	.73	.91	.73	.61	.44	.48	- .85	-1.01	-3.01	-2.76	1.08	.78	.49	.55
	8	284	.78	.88	.86	.77	.78	.70	- .91	-1.02	-1.52	-1.69	1.72	1.22	1.24	.99
	9	285	.74	.86	.54	.61	.67	.48	-1.22	-1.07	-1.63	-2.28	.63	.77	.90	.55
	10	286	.86	.96	.74	.57	.89	.74	-1.47	-1.92	-1.95	-2.34	1.10	.69	1.97	1.11
	11	287	.36	.35	.37	.12	.31	.22	.99	2.91	1.27	1.75	.39	.13	.33	.23
	12	288	.35	.43	.49	.40	.73	.48	.77	.93	.23	.35	.56	.44	1.06	.55

(continued)

Table 7 (continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
25 Map Directions	1	289	.74	.88	.55	.60	.56	.48	-1.17	-1.06	-2.06	-2.43	.66	.75	.68	.54
	2	290	.50	.54	.45	.29	.43	.30	.01	.02	- .24	- .34	.50	.30	.48	.31
	3	291	.72	.85	.66	.60	.65	.53	- .90	- .98	-1.57	-1.92	.88	.76	.86	.63
	4	292	.83	.94	.60	.51	.63	.49	-1.59	-1.89	-2.51	-3.25	.76	.59	.82	.56
	5	293	.81	.84	.45	.56	.70	.65	-1.91	-1.54	-1.45	-1.57	.51	.67	.99	.86
	6	294	.29	.24	.50	.28	.53	.27	1.09	1.92	1.30	2.51	.58	.30	.63	.29
	7	295	.29	.15	.30	.16	.50	.41	1.85	3.62	2.07	2.52	.32	.16	.57	.44
	8	296	.52	.69	.60	.61	.67	.68	- .07	- .07	- .76	- .75	.76	.76	.90	.92
	9	297	.27	.31	.42	.26	.63	.26	1.50	2.40	.81	1.98	.46	.27	.80	.27
	10	298	.61	.71	.47	.33	.61	.49	- .60	- .85	- .92	-1.17	.53	.35	.79	.56
	11	299	.52	.61	.54	.48	.52	.36	- .08	- .09	- .55	- .80	.64	.55	.61	.38
	12	300	.40	.42	.55	.41	.64	.47	.46	.62	.30	.41	.66	.45	.84	.54
26 Map Measurement	1	301	.70	.87	.58	.57	.70	.75	- .91	- .94	-1.62	-1.51	.72	.69	.99	1.14
	2	302	.68	.73	.52	.48	.63	.63	- .90	- .98	- .99	-1.00	.62	.55	.82	.80
	3	303	.64	.81	.68	.65	.31	.22	- .51	- .54	-2.81	-4.03	.93	.85	.33	.22
	4	304	.69	.81	.47	.66	.69	.69	-1.03	- .74	-1.27	-1.28	.54	.89	.97	.95
	5	305	.54	.54	.65	.59	.60	.40	- .15	- .16	- .17	- .26	.85	.73	.76	.44
	6	306	.32	.32	.52	.50	.52	.32	.91	.95	.92	1.51	.61	.58	.61	.33
	7	307	.35	.37	.54	.35	.39	.30	.72	1.11	.84	1.08	.64	.37	.42	.32
	8	308	.48	.63	.62	.60	.67	.48	.07	.08	- .50	- .70	.78	.75	.91	.55
	9	309	.51	.59	.65	.58	.67	.64	- .05	- .06	- .33	- .34	.86	.71	.89	.83
	10	310	.35	.40	.25	.22	.46	.50	1.49	1.70	.56	.52	.26	.23	.52	.58
	11	311	.54	.67	.65	.46	.49	.54	- .17	- .24	- .93	- .83	.85	.52	.56	.64
	12	312	.39	.41	.26	.13	.39	.26	1.02	1.98	.56	-2.22	.27	.14	.43	.27
27 Map Scale	1	313	.62	.50	.52	.39	.56	.43	- .59	- .80	.00	.00	.62	.42	.68	.48
	2	314	.68	.73	.56	.69	.65	.72	- .81	- .67	- .96	- .87	.68	.95	.86	1.04
	3	315	.81	.86	.53	.57	.66	.60	-1.63	-1.50	-1.63	-1.79	.62	.70	.87	.74
	4	316	.75	.88	.62	.57	.55	.57	-1.11	-1.20	-2.11	-2.03	.79	.70	.66	.70
	5	317	.69	.82	.57	.50	.75	.73	- .87	-1.00	-1.19	-1.23	.70	.58	1.15	1.07
	6	318	.43	.48	.52	.41	.51	.44	.36	.46	.10	.12	.60	.45	.59	.48
	7	319	.43	.40	.52	.44	.54	.44	.34	.40	.48	.59	.60	.48	.64	.49
	8	320	.50	.69	.50	.34	.76	.67	- .01	- .02	- .65	- .74	.58	.36	1.16	.90
	9	321	.46	.67	.53	.47	.61	.67	.18	.21	- .71	- .65	.63	.53	.77	.91
	10	322	.40	.48	.56	.49	.49	.47	.45	.51	.08	.08	.68	.57	.56	.53
	11	323	.64	.65	.63	.48	.48	.33	- .57	- .75	- .82	-1.19	.81	.55	.55	.35
	12	324	.36	.36	.57	.58	.68	.63	.63	.62	.52	.56	.69	.71	.93	.82
28 Longitude	1	325	.87	.91	.67	.74	.55	.65	-1.67	-1.51	-2.46	-2.08	.89	1.09	.67	.86
	2	326	.50	.52	.51	.46	.73	.55	.01	.01	- .07	- .09	.60	.52	1.08	.66
	3	327	.67	.67	.58	.55	.73	.67	- .74	- .78	- .62	- .67	.72	.66	1.06	.91
	4	328	.79	.80	.70	.65	.88	.63	-1.17	-1.27	- .97	-1.33	.98	.86	1.81	.82
	5	329	.65	.81	.57	.53	.88	.75	- .66	- .71	- .98	-1.16	.69	.62	1.85	1.12
	6	330	.46	.37	.38	.08	.28	.04	.25	1.16	1.14		.41	.08	.30	.04
	7	331	.32	.31	.36	.20	.30	.03	1.30	2.32	1.67		.39	.21	.31	.03
	8	332	.73	.84	.78	.70	.73	.80	- .80	- .89	-1.38	-1.25	1.25	.98	1.06	1.33
	9	333	.55	.72	.75	.72	.81	.72	- .18	- .19	- .72	- .81	1.15	1.04	1.36	1.03
	10	334	.55	.69	.67	.62	.73	.67	- .20	- .22	- .69	- .76	.91	.79	1.08	.91
	11	335	.23	.26	.16	.22	.37	.39	4.70	3.38	1.72	1.63	.16	.24	.40	.43
	12	336	.31	.38	.67	.60	.56	.51	.72	.82	.53	.58	.91	.74	.67	.59

(continued)

Table 7 (continued)

Concept	Task	Item	Proportion Correct		Biserial Correlation				X_{50}				Beta			
					Boys[a]		Girls[b]		Boys		Girls		Boys		Girls	
			Boys	Girls	C	T	C	T	C	T	C	T	C	T	C	T
29 Physical Feature Map	1	337	.75	.90	.68	.79	.59	.64	-1.01	- .87	-2.15	-1.98	.92	1.30	.73	.84
	2	338	.67	.71	.75	.72	.56	.50	- .59	- .62	-1.02	-1.13	1.15	1.05	.67	.58
	3	339	.56	.53	.63	.50	.70	.64	- .23	- .29	- .11	- .12	.82	.58	.99	.83
	4	340	.83	.83	.59	.52	.55	.47	-1.59	-1.81	-1.71	-2.00	.73	.61	.66	.53
	5	341	.67	.71	.64	.58	.80	.71	- .70	- .77	- .71	- .80	.83	.71	1.31	1.00
	6	342	.51	.63	.75	.69	.64	.59	- .04	- .05	- .51	- .56	1.13	.94	.82	.72
	7	343	.31		.56	.55			.90	.92			.67	.66		
	8	344	.45	.59	.77	.65	.80	.73	.16	.19	- .29	- .32	1.22	.86	1.34	1.07
	9	345	.47	.51	.67	.62	.65	.54	.11	.11	- .04	- .05	.91	.80	.86	.64
	10	346	.41	.47	.55	.40	.66	.44	.41	.57	.12	.17	.66	.44	.87	.49
	11	347	.35	.37	.43	.31	.67	.57	.90	1.24	.51	.59	.48	.33	.90	.70
	12	348	.30	.30	.66	.69	.57	.61	.81	.77	.94	.88	.88	.96	.69	.78
30 Symbol Map	1	349	.88	.91	.70	.82	.56	.78	-1.68	-1.45	-2.44	-1.74	.99	1.41	.67	1.25
	2	350	.73	.80	.64	.50	.78	.71	- .95	-1.20	-1.[illegible]	-1.16	.83	.59	1.25	1.01
	3	351	.78	.90	.71	.71	.60	.57	-1.12	-1.11	-2.16	-2.29	1.00	1.00	.75	.69
	4	352	.75	.83	.62	.55	.67	.42	-1.09	-1.21	-1.41	-2.23	.79	.66	.90	.47
	5	353	.72	.90	.75	.74	.78	.76	- .76	- .78	-1.66	-1.70	1.15	1.11	1.26	1.18
	6	354	.43	.60	.44	.39	.39	.29	.43	.48	- .62	- .85	.49	.42	.43	.30
	7	355	.44	.39	.47	.41	.56	.53	.32	.36	.48	.51	.53	.45	.68	.63
	8	356	.46	.64	.62	.59	.69	.55	.16	.16	- .51	- .64	.79	.72	.94	.66
	9	357	.67	.85	.65	.61	.70	.49	- .66	- .70	-1.47	-2.10	.86	.77	.97	.56
	10	358	.35	.30	.36	.32	.49	.37	1.09	1.22	1.06	1.40	.38	.34	.56	.40
	11	359	.51	.72	.74	.63	.70	.67	- .03	- .03	- .85	- .89	1.10	.82	.98	.90
	12	360	.33	.32	.59	.46	.58	.44	.76	.96	.80	1.06	.72	.52	.71	.49

[a] N = 195

[b] N = 196

correlation for these two cases was greater than 1.00, and the X_{50} and β could not be computed. The β for these items can be interpreted as being infinity.

If desired, the items that make up a criterion score can be separated out. This is easy to do for a concept; the items composing the criterion score are simply the 12 given in order consisting of one of each task type. For example, the items composing the criterion score for Concept 3 are numbered 25 through 36. The items composing the criterion score for a task are those with the same task number for each of the concepts; for example, the items composing the criterion score for Task 1 are numbered 1, 13, 25, 37, etc., with the last one being number 349.

As was evident from the means of the total scores, and as can be seen from the two difficulty indices given for the items (proportion correct and X_{50}), the items, in general, were more difficult for the boys than for the girls. There is not a one-to-one correspondence for each item, however; there are some exceptions, since some items were more difficult for the girls and some were about the same. As was pointed out earlier, however, no conclusions can be drawn from this because the data for the girls were collected in early summer shortly after the end of their fifth grade school year and the data for the boys were collected in the fall shortly after their sixth grade school year had begun. The difficulty indices obtained indicate that these items are of appropriate difficulty levels for these subjects.

It seems clear from looking at Table 7

that X_{50} gives more precise information about the difficulty level of an item when that same item is a part of each of two criterion scores. The proportion correct remains the same for both of the criterion scores. This index tells how many subjects responded to the correct answer for an item but it says nothing about their ability level as measured by a particular criterion score—total concept score or total task score in this case. The item difficulty index, X_{50}, gives (in standard deviation units) the criterion score at which a subject would have a 50-50 chance of getting the item correct. For example, an X_{50} value of 1.20 for an item indicates that subjects with a criterion score 1.20 standard deviation units *above* the mean have a 50% chance of answering that item correctly. Subjects with a criterion score higher than this would have a greater chance of answering that item correctly, and subjects with a criterion score lower than this would have a lesser chance. Likewise, an X_{50} value of -1.20 means that subjects with a criterion score 1.20 standard deviations *below* the mean would have a 50% chance of getting that item correct; for a higher score the chance would be greater, and for a lower score the chance would be less. Knowing both X_{50} and β for an item allows one to readily determine the probability of answering an item correctly for any point on the criterion scale (Baker, 1964). It may be pointed out that when P = .50, X_{50} = .00; when P is greater than .50 then X_{50} will be negative and, for a certain P, the higher the β value the closer to zero will be the X_{50} value. This can be seen from inspecting Table 7. For example, for Item 1 the β is higher for the concept score than it is for the task score for both boys and girls; similarly for both boys and girls, the X_{50} value is closer to zero for the concept score than it is for the task score. For P less than .50, the X_{50} will be positive, and again, for a certain P, the higher the β value the closer to zero will be the X_{50} value. See Item 347 for an illustration of this.

The two item discrimination indices, biserial correlation and β, are more closely related since β is computed as a function of the biserial correlation (Baker, 1969). They are not linearly related, however. From .00 to about .30 (absolute) they are very nearly the same; beyond this, β begins to increase quite rapidly in magnitude. It may be pointed out that β is always equal to or greater (absolute) than the biserial correlation. As a general rule, .30 is often used as a lower cutting point for a desirable biserial correlation or β. For a total score composed of relatively few items, as is the concept score, a much higher minimum would be desirable.

As can be seen from Table 7, almost all of the social studies items have desirable biserial correlations and βs when the item is both a part of a concept criterion score and a task criterion score, and most of them are very good. The most obvious thing is that the βs are higher, with a few exceptions, when the item is a part of a concept criterion score than when it is a part of a task criterion score. This is to be expected since the concept score consists of considerably fewer items than does the task score—12 items for most concept scores and 30 items for most task scores. The item-criterion biserial correlation is a part-whole correlation, with the criterion the total score of which the item is a part, and the fewer the number of items the greater should be the correlation of that item with the total score of which it is a part. Since β is computed as a function of the biserial correlation, it is affected in the same manner. There does not seem to be a consistent pattern in the magnitude of the βs for the boys as compared with the girls. For some of the items, the βs are considerably higher for the boys, and for some of them they are considerably higher for the girls. For the tryouts of the items, data for both boys and girls were analyzed together. If the data for boys and girls were pooled and item analyzed, the β values would probably increase for most of the items.

As was discussed earlier, these item indices were obtained by performing conventional item analyses on two different types of scores—one for concept criterion scores and one for task criterion scores. This was necessitated by the lack of item analysis procedures appropriate for use with data collected using a completely crossed design to build the items. It is not known how the item indices would be affected if procedures were available to compute them simultaneously taking into account the effects of the crossed design. A guess would be that discrimination indices would be affected more than would difficulty indices, if there were an effect. It it plausible to expect that there may be some concept-task interactions which cannot be, at least readily, ascertained by doing a conventional item analysis on the two types of scores.

IV
Summary and Conclusions

The primary objective of the project entitled "A Structure of Concept Attainment Abilities" is to formulate one or more models or structures of concept attainment abilities, and to assess their consistency with actual data. One of the major steps for attaining this primary objective was taken to be the development of tests to measure achievement of selected language arts, mathematics, science, and social studies concepts appropriate at the fourth grade level. This paper describes the test development efforts and presents the item and total score statistics obtained using the revised items developed for measuring achievement of selected concepts in social studies.

Subject matter specialists identified single or compound word classificatory concepts for three major areas, and randomly selected 10 from each area to be studied. These 30 selected concepts were then analyzed. Twelve items for each concept were developed; one for each of the first 12 tasks of "A Schema for Testing the Level of Concept Mastery" (Frayer, Fredrick, & Klausmeier, 1969).

The items that were developed were administered during early summer of 1970 to 196 girls who had just completed the fifth grade and during the fall of 1970 to 195 boys who had just begun the sixth grade. These data were item analyzed, separately for boys and for girls, using the GITAP program (Baker, 1969).

The means, standard deviations, and Hoyt reliability estimates obtained are presented and discussed for total concept and total task scores. Four different item indices—proportion correct, item-criterion biserial correlation, X_{50}, and β—obtained for each item based on each of two criterion scores, appropriate total concept score and appropriate total task score, are presented and discussed.

Conclusions

The major conclusions drawn are:

1. The reliability estimates obtained for both total concept scores and total task scores are sufficiently high to warrant study of the dimensionality of these selected social studies concepts and the dimensionality of the tasks when using social studies content.
2. The three area distinctions seem not to be important ones.
3. The difficulty item indices obtained indicate that these items are of appropriate difficulty levels for these subjects.
4. Almost all of the items have desirable levels of discrimination indices when the item is both a part of a concept criterion score and a task criterion score.

Recommendation

The completely crossed design used to construct these achievement tests is a very interesting one. This type of design might well be used more often in the future. It would be highly desirable to have available item analysis procedures that are appropriate for analyzing such crossed designs. At the present such a methodology is not known.

GPO 827-034-3

References

Baker, F. B. An intersection of test score interpretation and item analysis. *Journal of Educational Measurement*, 1964, *1*, 23-28.

Baker, F. B. FORTAP: a fortran test analysis package. Department of Educational Psychology, The University of Wisconsin, 1969.

Bourne, L. E., Jr. *Human conceptual behavior*. Boston: Allyn & Bacon, 1966.

Frayer, D. A., Fredrick, W. C., & Klausmeier, H. J. *A schema for testing the level of concept mastery*. Wisconsin Research and Development Center for Cognitive Learning, Working Paper No. 16, 1969.

Harris, M. L., Harris, C. W., Frayer, D. A., & Quilling, M. R. *A structure of concept attainment abilities: The problem and strategies for attacking it*. Wisconsin Research and Development Center for Cognitive Learning, Theoretical Paper No. 32, in press.

Klausmeier, H. J., Harris, C. W., Davis, J. K., Schwenn, E., & Frayer, D. *Strategies and cognitive processes in concept learning*. The University of Wisconsin, Cooperative Research Project No. 2850, 1968.

Tabachnick, B. R., Weible, E. B., & Frayer, D. A. *Selection and analysis of social studies concepts for inclusion in tests of concept attainment*. Wisconsin Research and Development Center for Cognitive Learning, Working Paper No. 53, 1970.

Tabachnick, B. R., Weibel, E. B., & Livermore, D. *Items to test level of attainment of social studies concepts by intermediate grade children*. Wisconsin Research and Development Center for Cognitive Learning, Working Paper No. 54, 1970.

Appendix A
Key for Concepts in Numerical Order

1 Coastline
2 Delta
3 Desert
4 Gulf
5 Mountain Region
6 River
7 River Mouth
8 Strait
9 Tributary
10 Tropical Region
11 Airway
12 City
13 Countryside
14 Democracy
15 Exchange
16 Government
17 Land Routes
18 News
19 Organization
20 Waterway
21 Country
22 Distance
23 East-West Lines of Latitude
24 Globe
25 Map Directions
26 Map Measurement
27 Map Scale
28 North-South Lines of Longitude
29 Physical Feature Map
30 Symbol Map

GPO 827-034-2

National Evaluation Committee

Helen Bain
Immediate Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
Universiy of Colorado

Jeanne S. Chall
Graduate School of Education
Harvard University

Francis S. Chase
Department of Education
University of Chicago

George E. Dickson
College of Education
University of Toledo

Hugh J. Scott
Superintendent of Public Schools
District of Columbia

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Benton J. Underwood
Department of Psychology
Northwestern University

Robert J. Wisner
Mathematics Department
New Mexico State University

Executive Committee

William R. Bush
Director of Program Planning and Management
and Deputy Director, R & D Center

Herbert J. Klausmeier, Committee Chairman
Director, R & D Center

Wayne Otto
Principal Investigator
R & D Center

Robert G. Petzold
Professor of Music
University of Wisconsin

Richard A. Rossmiller
Professor of Educational Administration
University of Wisconsin

James E. Walter
Coordinator of Program Planning
R & D Center

Russell S. Way, ex officio
Program Administrator, Title III ESEA
Wisconsin Department of Public Instruction

Faculty of Principal Investigators

Vernon L. Allen
Professor of Psychology

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
Center Director
V. A. C. Henmon Professor
Educational Psychology

Stephen J. Knezevich
Professor
Educational Administration

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

Wayne Otto
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Peter A. Schreiber
Assistant Professor
English

Richard L. Venezky
Associate Professor
Computer Science

Alan M. Voelker
Assistant Professor
Curriculum and Instruction

Larry M. Wilder
Assistant Professor
Communication Arts